



Technisches Handbuch openTA-Newsdienst

openTA-Dokument Nr. 3

Stand: 11.06.2014, Version 1.1

Autoren: Christian Schmitt, Rainer Weidemann (KIT-IAI)

Dieses Dokument beschreibt den openTA-Newsdienst und richtet sich in erster Linie an die technisch verantwortlichen Personen der jeweiligen NTA-Mitgliedsinstitutionen. Der openTA-Newsdienst wird durch zwei weitere Dokumente ergänzend beschrieben:

- openTA: openTA-Newsdienst – Formale Kategorisierung (openTA-Dokument Nr. 1)
- openTA: openTA-Newsdienst – Ziele, Verfahren, Nutzung (openTA-Dokument Nr. 2)

Die Implementierung des openTA-Newsdienst entspricht teilweise noch nicht dem hier beschriebenen Stand. *Darauf wird jeweils hingewiesen.*

Kontakt: info@openTA.net

Die openTA-Dokumente sind abrufbar unter der Adresse
<http://www.openta.net/veroeffentlichungen-und-openta-dokumente>

Versionsgeschichte:

Mai 2013, erster Entwurf

Januar 2014 Version 1.0 (redaktionelle Änderungen und Aktualisierungen)

Juni 2014 Version 1.1 (Update technischer Details an Implementierungsfortschritt)

Rechtehinweis: CC BY-SA 4.0

Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International

Der openTA-Newsdienst ist Teil des Forschungsvorhabens „Kooperativer Aufbau eines Fachportals Technikfolgenabschätzung“, gefördert durch die DFG, durchgeführt von den KIT-Instituten ITAS, IAI und der Bibliothek.

Inhaltsverzeichnis

1	Vorwort.....	5
2	Anforderungen an aggregierbare News-Feeds	5
2.1	Inhaltselemente auf Feed-Ebene (oberste Ebene)	6
2.2	Inhaltselemente auf Beitragsebene (Entry bzw. Item).....	7
2.3	Copyright- bzw. Lizenz-Angabe.....	9
2.4	Erweiterte Inhalte.....	10
	Raumbezug	10
	Medien	10
3	Bereitstellung von Beiträgen über Microformate in HTML-Seiten.....	11
4	Eingabe von Beiträgen im openTA-Portal über ein Webformular.....	11
5	Nutzung des openTA-Newsdienstes.....	12
5.1	Nachrichtenticker im openTA-Portal.....	12
5.2	Konfigurierbarer Atom-Newsfeed	13
5.3	Recherche im News-Archiv.....	13
5.4	Web-Widget.....	14
5.5	E-Mail-Dienst	16
6	Literatur/Informationsquellen.....	17
7	Anhang A: Beispiel-Feed mit openTA-Kategorien im Atom-Format.....	18
8	Anhang B: Beispiel-Feed mit openTA-Kategorien im RSS-Format	19

Tabellenverzeichnis

Tab. 1:	Inhaltselemente für die Feed-Ebene: openTA-Anforderungen (Atom/RSS)	6
Tab. 2:	Inhaltselemente auf der Beitragsebene: openTA-Anforderungen (Atom/RSS)	7
Tab. 3:	Angabe eines Raumbezugs per GeorSS-simple-Format.....	10
Tab. 4:	Angabe von Medien (Bilder, Video, Audo) per MediaRSS	11
Tab. 5:	Vorgesehene Felder für ein Webformular für die Eingabe von Nachrichten.....	12
Tab. 6:	Konfigurationsparameter für den openTA-Newsfeed.....	13
Tab. 7:	Rechercheparameter für die Suche im openTA-News-Archiv	14
Tab. 8:	Datenattribute für das News-Widget.....	15

Abbildungsverzeichnis

Abb. 1:	Create Commons Lizenz.....	10
Abb. 2:	openTA-News-Widget selbst einbinden.....	16

1 Vorwort

Im Rahmen des DFG-Projektes openTA wird ein Dienst entwickelt, welcher Neuigkeiten aus der TA-Community aufammelt, aufbereitet und integriert bereitstellt. Im Folgenden (Kap.2) wird beschrieben, wie die eigenen News-Feeds der NTA-Mitgliedsinstitutionen aufgebaut sein müssen, um von openTA aggregiert werden zu können. Beiträge können auch in speziellen, mit Mikroformaten ausgezeichneten HTML-Seiten oder „per Hand“, d.h. über ein Formular im openTA-Portal, eingegeben werden, was Thema der nächsten beiden Kapitel 3 und 4 ist. Schließlich wird in Kapitel 5 beschrieben, wie der openTA-Newsdienst mit den aggregierten Informationen von den beteiligten NTA-Institutionen und von Dritten genutzt werden kann.

2 Anforderungen an aggregierbare News-Feeds

Der bevorzugte Weg zur Bereitstellung von aktuellen, in openTA aggregierbaren Informationen sind News-Feeds. Diese müssen über eine eindeutige und permanente Adresse (URL) über das Internet erreichbar und ohne Autorisierung zugreifbar sein. Der openTA-Server wird diese News-Feeds in regelmäßigen Intervallen abfragen, zur Zeit erfolgt diese Abfrage täglich um 12 Uhr. Nach einer Überprüfung auf technische Kompatibilität zum openTA-Newsdienst wird der Feed dann in den News-Dienst aufgenommen.

In einer Ausbaustufe des Mitgliederportals sollen Vorgaben für die Abfrageintervallen und weitere Angaben zur Kennzeichnung und Verarbeitung der Feeds der Mitgliedsinstitutionen in der Dienstanmeldung (Teil der Mitgliederverwaltung für institutionelle Mitglieder) eingetragen werden.

Als Feed-Format wird Atom empfohlen und im openTA-Newsdienst intern verwendet, da dieses das modernere, XML-Namespaces-kompatible und beliebig erweiterbare Newsformat ist. openTA unterstützt aber auch Feeds im Format RSS 2.0, beschränkt auf die mit Atom kompatiblen Inhaltselemente. Ältere RSS-Formate werden nicht explizit unterstützt, können aber angebunden werden, wobei davon auszugehen ist, dass nur die wichtigsten, mit Atom und RSS 2.0 kompatiblen Inhaltselemente weitergenutzt werden.

Jeder Feed, der in den openTA-Newsdienst eingebunden wird, wird organisatorisch einer NTA-Institution zugeordnet. Die Zuordnung wird bei der Anmeldung des Feeds durch den Verantwortlichen der NTA-Mitgliedsinstitution festgelegt.¹ Angaben zu „author“ und „contributor“ (s.u.) im Feed werden als Zusatzinformationen zum aktuellen Inhalt betrachtet.

Einzelbeiträge im Feed können in verschiedenen Sprachen geschrieben sein (normalerweise deutsch oder englisch). Jeder Einzelbeitrag wird durch den openTA-Newsdienst genau einer Sprache zugeordnet. Die Zuordnung erfolgt bei RSS über das Feld „language“ einheitlich für alle Beiträge eines Feeds. Bei Atom-Feeds erfolgt die Zuordnung individuell für jeden Beitrag über den Wert des Attributs „xml:lang“ beim Entry-Tag des Einzelbeitrags. Standard-Sprache (keine Angabe) ist in beiden Formaten deutsch.

Die folgende Tabelle enthält die vom openTA-Newsdienst verarbeitbaren Inhaltselemente (Felder) von News-Feeds und beschreibt, welche Inhalte in diesen Feldern erwartet wer-

¹ Die Anmeldung erfolgt über info@openTA.net. In einer späteren Ausbaustufe des Mitgliederportals soll sie in die dort vorgesehene Dienstanmeldung integriert werden.

den. Weitere, laut der jeweiligen Spezifikation mögliche Felder werden von openTA ignoriert. Zusätzlich sind in der Tabelle die Felder gekennzeichnet, die unbedingt benötigt werden (P = Pflicht) und die für inhaltlich vollständige Beiträge sowie für eine übergreifende Kategorisierung empfohlen werden (E = erwünscht/empfohlen).

2.1 Inhaltselemente auf Feed-Ebene (oberste Ebene)

Inhaltselemente auf Feed-Ebene² beschreiben den Rahmen, in den die Einzelbeiträge eingebettet sind. Da der openTA-Newsdienst die Einzelbeiträge aus diesem Rahmen herausholt und als aggregierten Feed mit eigenen Daten neu verpackt, sind einige Inhaltselemente der Feed-Ebene für den openTA-Newsdienst ohne Bedeutung. Ausgewertet werden nur die, welche als Vorbelegungen für Inhaltselement der eingeschlossenen Einzelbeiträge verwendet werden können, falls zu diesen keine spezifischen Werte angegeben sind.

Tab. 1: Inhaltselemente für die Feed-Ebene: openTA-Anforderungen (Atom/RSS)

Feld / Tag		Status	Inhalt
Atom	RSS 2.0		
title	-	O, S	Bezeichnung des Feeds Format: Einfacher Text, HTML- oder XHTML-Text (Atom: type="text html xhtml", Default = "text") Beispiel: <title>Neues aus ITAS</title>
updated	last-BuildDate	O, S	Datum der letzten Änderung des Feeds Die Angabe wird nur herangezogen, wenn Einzelbeiträge keine Datumsangaben haben (s.u.) Format: Datumsangabe nach [1] bzw. [2] Beispiele: Atom: <updated>2013-06-03T10:55:21.000Z</updated> RSS: <lastBuildDate> Mon, 03 Jun 2013 10:55:21 +0000</lastBuildDate>
-	pubDate	O, S	Veröffentlichungsdatum: des Feeds Die Angabe wird nur herangezogen, wenn Einzelbeiträge keine Datumsangaben haben (s.u.) Format: Datumsangabe nach [2]
-	language	E, S	Sprache (kann in RSS nur auf Feed-Ebene angegeben werden, d.h. es werden getrennte Feeds für jede Sprachvariante empfohlen) Format: Kürzel nach [4] Default: de Beispiel (RSS): <language>en</language>

² Das Handbuch ist keine Einführung in Atom und RSS und geht daher davon aus, dass die grundsätzliche Struktur eines Atom- bzw. RSS-Feed und die Bedeutung der Inhaltselemente bekannt sind. Siehe dazu z.B. [1], [2]

author	-	O,M	Autor (Person oder Institution): Die organisatorische Zuordnung eines Feeds zu einer NTA-Institution erfolgt automatisch bei der Einbindung des Feeds in den openTA-Newsdienst. Die Angaben zu weiteren Autoren auf Feed-Ebene werden als Zusatzinformationen betrachtet, die zusätzlich neben der ohnehin vorhandenen Zuordnung zu einer NTA-Institution bei jedem Beitrag aufgeführt werden sollen. Format: Dieses Inhaltselement besitzt eine Unterstruktur (siehe nächste Tabellenzeilen). Falls ein Autor angegeben wird, muss dieser einen Namen, sollte eine E-Mail-Adresse und kann eine URI erhalten. Beispiel (mit Name und Email-Adresse): <author> <name>Frank N. Stein</name> <email>frank.n.stein@example.com</email> </author>
author/name	-	P,S	Name eines Autors (als Unterelement zu „author“): Format: Einfacher Text, HTML- oder XHTML-Text (Atom: type=“text html xhtml“, Default = “text“)
author/email	-	E,S	Email-Adresse des Autors (als Unterelement zu „author“): Format: Valide Email-Adresse
author/uri	-	O,S	Home-Page des Autors (als Unterelement zu „author“): Format: Valide URI
contributor	-	O,M	Mitwirkende (Person oder Institution): Ergänzt zusätzlich nochmal den oder die Autoren für alle Einzelbeiträge des Feeds. Format: Das Inhaltselement besitzt dieselbe Unterstruktur wie ein Autor (hier nicht einzeln ausgeführt)
category	category	O,M	Kategorie: Kategorien sollten besser auf Beitragsebene angegeben werden (s.u.). Für Feeds angegebene Kategorien werden in openTA allen Beiträgen zugeordnet, soweit nicht auf der Beitragsebene entsprechende Kategorien vorhanden sind. Für einen monokategorialen Feed (etwa Publikationsankündigungen) würde die Angabe der Kategorie auf Feed-Ebene reichen. Format: wie auf Beitragsebene (s.u.)
rights	copyright	E,S	Copyright für den Feed insgesamt, d.h. für alle Beiträge, zu denen es keine spezifische Angabe (s.u.) gibt. Bei RSS kann das Copyright nur auf Feed-Ebene definiert werden. Format: Atom: Einfacher Text, HTML- oder XHTML-Text (type=“text html xhtml“, Default = “text“) RSS: Einfacher Text weitere Details und Beispiele siehe 2.3

Status: P = Pflicht, E = erwünscht, O = optional (in Bezug auf das übergeordnete Element);
S = nur einmal pro übergeordnetem Element, M = kann mehrmals angegeben werden
Vorgabe für openTA kann von der Spezifikation des Formats abweichen

Ignoriert werden folgende Felder (Atom): id, subtitle, link, icon³, logo, generator.

2.2 Inhaltselemente auf Beitragsebene (Entry bzw. Item)

Tab. 2: Inhaltselemente auf der Beitragsebene: openTA-Anforderungen (Atom/RSS)

Feld / Tag		Status	Inhalt
Atom	RSS 2.0		
entry, Attribut xml:lang	-	E,S	Attribut: Sprache des Beitrags, sollte bei fremdsprachigen Beiträgen unbedingt angegeben werden Format: Kürzel nach [4] als Attributwert; Default: de, falls keine Vorgabe für den Feed im Mitgliederbereich hinterlegt ist. Beispiel: <entry xml:lang=“en“> ... </entry>
id	guid	P,S	Eindeutiger (zumindest bezogen auf den jeweiligen Feed, besser global) und permanenter Bezeichner des Beitrags Format: Text; empfohlen wird die Angabe einer URI Beispiel: <id>http://example.com/ausschreibung/4711</id>

³ Logos der NTA-Institutionen werden in der openTA-Mitgliederverwaltung hinterlegt und von dort aus an passender Stelle eingefügt. Im Feed angegebene Institutslogos werden ignoriert.

⁴ Im ersten Ansatz erfolgt eine formale Kategorisierung nach der Art des Beitrags [3]. Diskutiert

title	title	P,S	<p>Titel des Beitrags</p> <p>Format:</p> <p>Atom: Einfacher Text, HTML- oder XHTML-Text (type="text html xhtml", Default = "text")</p> <p>RSS: Einfacher Text</p> <p>Beispiel: <title>Stellenausschreibung Techniker</title></p>
link, Attribut rel = "alternate"	link	E,S	<p>Verweis auf eine Web-Seite mit dem vollständigen Beitrag (zur Verlinkung auf den Originalbeitrag bei NTA-Institution erforderlich)</p> <p>Format</p> <p>Siehe [1] bzw. [2]</p> <p>Beispiel:</p> <p>Atom: <link rel="alternate" href="http://example.com/ausschreibung/4711" /></p> <p>RSS: : <link>http://example.com/ausschreibung/4711</link></p>
link, Attribut rel="enclosure"	enclosure	O, M	<p>Verweis auf Multimedia-Inhalte (bspw. MP3-Datei)</p> <p>Format</p> <p>Siehe [1] bzw. [2]</p> <p>Beispiel:</p> <p>Atom: <link rel="enclosure" href="http://example.com/podcast/0815.mp3" /></p> <p>RSS: <enclosure>http://example.com/podcast/0815.mp3</enclosure></p>
summary	description	E,S	<p>Kurze Zusammenfassung, Text-Snippet (sollte unbedingt enthalten sein)</p> <p>Format:</p> <p>Atom: Einfacher Text, HTML- oder XHTML-Text (type="text html xhtml", Default = "text")</p> <p>RSS: Einfacher Text oder entity-encoded HTML</p>
content	(content)	E,S	<p>Vollständiger Beitrag, sollte unbedingt vorhanden sein, wenn kein Link (s.o.) angegeben ist, ansonsten optional aber empfohlen.</p> <p>Für RSS ist eigentlich kein content-Element vorgesehen, es wird von openTA jedoch unterstützt (via RSS-Erweiterung http://web.resource.org/rss/1.0/modules/content/)</p> <p>Format und Beispiele:</p> <p>Atom: Einfacher Text, HTML- oder XHTML-Text (type="text html xhtml", Default = "text")</p> <p>RSS: Einfacher Text oder entity-encoded HTML</p>
published	-	O,S	<p>Veröffentlichungsdatum des Beitrags (1. Version)</p> <p>Format:</p> <p>Datumsangabe nach [1] oder [2]</p>
updated	pubDate	P,S	<p>Datum der letzten substantiellen Änderung des Beitrags (identisch zu published bei 1. Version, späteres Datum nach Änderungen)</p> <p>Format:</p> <p>Datumsangabe nach [1]</p>
author	-	O, M	<p>Autor (Person oder Institution) des Beitrags:</p> <p>Die organisatorische Zuordnung eines Beitrags zu einer NTA-Institution erfolgt automatisch bei der Einbindung des Feeds in den openTA-Newsdienst. Die Angaben zu weiteren Autoren zum Beitrag werden als Zusatzinformationen betrachtet, die zusätzlich neben der ohnehin vorhandenen Zuordnung zu einer NTA-Institution bei dem Beitrag aufgeführt werden sollen.</p> <p>Format:</p> <p>siehe oben auf Feed-Ebene</p>
author/name	-	P,S	<p>Name eines Autors (als Unterelement zu „author“):</p> <p>Format:</p> <p>Einfacher Text, HTML- oder XHTML-Text (Atom: type="text html xhtml", Default = "text")</p>
author/email	author	E,S	<p>Email-Adresse des Autors (bei Atom als Unterelement zu „author“, bei RSS als eigenständiges Inhaltselement):</p> <p>Format:</p> <p>Valide Email-Adresse</p>
author/uri	-	O,S	<p>Home-Page des Autors (als Unterelement zu „author“):</p> <p>Format:</p> <p>Valide URI</p>
contributor	-	O, M	<p>Mitwirkende (Person oder Institution) des Beitrags:</p> <p>Ergänzt zusätzlich nochmal den oder die Autoren für den Einzelbeitrag.</p> <p>Format:</p> <p>Das Inhaltselement besitzt dieselbe Unterstruktur wie ein Autor (hier nicht einzeln ausgeführt)</p>
category	category	E, M	<p>Kategorie⁴ des Beitrags. Die Kategorien auf Beitragsebene ersetzen Kategorien höherer Ebenen (Feed, Mitgliederverwaltung)</p> <p>Format:</p> <p>siehe [3]; es können sowohl openTA-Kategorien als auch nicht openTA-Kategorien (Text) angegeben werden. Für die Filterung im openTA-Newsdienst werden jedoch nur die openTA-Kategorien herangezogen.</p>
rights	-	E,S	<p>Copyright für den Einzelbeitrag (ersetzt das Copyright des Feeds, falls definiert)</p> <p>Format:</p> <p>Atom: Einfacher Text, HTML- oder XHTML-Text (type="text html xhtml", Default = "text")</p> <p>weitere Details und Beispiele siehe 2.3</p>

⁴ Im ersten Ansatz erfolgt eine formale Kategorisierung nach der Art des Beitrags [3]. Diskutiert wurde auch eine weitere Kategorisierung nach Inhalten und Zielgruppen. Falls dafür ein Bedarf besteht, muss das Kategorienschema [3] entsprechend erweitert werden.

Status: P = Pflicht, E = erwünscht, O = optional (in Bezug auf das übergeordnete Element);
S = nur einmal pro übergeordnetem Element, M = kann mehrmals angegeben werden
Vorgabe für openTA kann von der Spezifikation des Formats abweichen

Ignoriert werden folgende Felder (Atom): source.

2.3 Copyright- bzw. Lizenz-Angabe

In den Feldern für die Angabe von Copyright- bzw. Lizenz-Informationen auf Feed- bzw. Beitragsebene können beliebige Informationen angegeben werden, welche auch unverändert übernommen werden, es wird jedoch die Angabe einer Creative-Commons-Lizenz [11] empfohlen (aktuelle Lizenzen: CC0 1.0, CC BY 3.0, CC BY-SA 3.0, CC BY-NC 3.0, CC BY-ND 3.0, CC BY-NC-SA 3.0, CC BY-NC-ND 3.0).

Die Angabe einer Creative-Commons-Lizenz setzt den Verweis auf die vollständigen Lizenz-Informationen voraus. Empfohlen wird hierfür der Verweis auf die entsprechende Seite unter creativecommons.org.

Beispiel: Vollständige Angabe für *CC BY-NC-SA 3.0 Unported*:

Atom:

```
<rights type="html">
  <a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">
    CC BY-NC-SA 3.0
  </a>
</rights>
```

RSS:

```
<copyright>
  CC BY-NC-SA 3.0, http://creativecommons.org/licenses/by-nc-sa/3.0/
</copyright>
```

Sollte diese ausführliche Angabe nicht möglich sein, erkennt der openTA-Crawler auch die Kürzel der Creative-Commons-Lizenzen und wandelt diese für die Ausgabe durch den openTA-News-Dienst in vollständige Angaben um. Ebenso ist es möglich, an die jeweilige Rechtssituation in einem speziellen Land angepasste Versionen zu verwenden.

Beispiel: Verkürzte Angabe für *CC BY-NC-SA 3.0 Deutschland*:

Atom: `<rights>CC BY-NC-SA 3.0 DE</rights>`

RSS: `<copyright>CC BY-NC-SA 3.0 DE</copyright>`

Das genaue Kürzel der verwendeten Lizenz ist auf der jeweiligen Lizenz-Seite auf creativecommons.org im Kopfbereich zu finden:

<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>



<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Abb. 1: Create Commons Lizenz

2.4 Erweiterte Inhalte

In den vorangegangenen Unterkapiteln wurde beschrieben, welche Informationen sich über die originären Strukturelemente des Atom- bzw. RSS-Formats beschreiben lassen. Sollen nicht darin enthaltene Informationskategorien eindeutig zuordenbar aufgenommen werden, müssen die Formate erweitert werden. Dies erfolgt über die Einbindung anderer, standardisierter XML-Formate als eigenständige XML-Namensräume [5], wie in dem folgenden Beispiel (Einbindung von GeORSS in Atom-Feed):

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom"
      xmlns:georss="http://www.georss.org/georss">
...
</feed>
```

Raumbezug

Die Beiträge eines Feeds sind einer NTA-Institution zugeordnet und werden bei einer räumlichen Suche über die Adresse der NTA-Institution verortet. Zusätzlich könnte zu einem Beitrag eine räumliche Zuordnung nach inhaltlichen Kriterien erfolgen, d.h. nicht um den Standort des Autors zu beschreiben, sondern den des im Beitrag behandelten Gegenstands (z.B. Ort einer Tagung). Dazu würde GeORSS [6] verwendet werden.

Der openTA-Newsdienst unterstützt nur die Angabe einer Punktkoordinate nach dem GeORSS-simple-Format. Eine Ortsangabe auf Feed-Ebene gilt für alle Einzelbeiträge, eine Angabe auf Beitragsebene nur für diesen Beitrag.

Tab. 3: Angabe eines Raumbezugs per GeORSS-simple-Format

Feld / Tag		Status	Inhalt
Atom	RSS 2.0		
georss:point	georss:point	O,S	WGS84-Koordinate (Wertepaar für Breite und Länge) Beispiel siehe [6] Beispiel: <georss:point>49.095 8.432</georss:point>

Medien

Als Alternative zu einem „Enclosure“ (s.o.) können in Newsbeiträgen Bilder, Videos, Audiodateien oder andere Medien-Dateien auch unter Verwendung von MediaRSS [9] eingebunden werden. Vorteil ist, dass hier ein zugehöriger Titel angegeben werden kann. Aktuell werden von openTA lediglich Vorschaubilder verarbeitet und wieder ausgegeben, weitere Medien-Typen werden ergänzt, wenn seitens der NTA-Mitgliedsinstitutionen Bedarf besteht.

Das Vorschaubild sollte sich auf den konkreten Inhalt des Einzelbeitrags beziehen und nicht dafür genutzt werden, ein allgemeines Organisationslogo zu transportieren. openTA fügt das Logo der NTA-Organisation automatisch ein, das hier angegebene Vorschaubild erscheint zusätzlich zum Logo.

Bilder können über den entsprechenden HTML-Tag (IMG) auch direkt in den Volltext des Newsbeitrags eingebunden werden, wenn dieser als HTML-Text erstellt wurde.

Tab. 4: Angabe von Medien (Bilder, Video, Audio) per MediaRSS

Feld / Tag		Status	Inhalt
Atom	RSS 2.0		
media:title	media:title	O,S	Einfacher Text Atom: Einfacher Text, HTML- oder XHTML-Text (type="text html xhtml", Default = "text") Beispiel: <media:title>Buchcover</media:title>
media:thumbnail	media:thumbnail	O,M	Attribut „url“ mit Vorschaubild-Adresse Beispiel: <media:thumbnail url="http://www.foo.de/cover.jpg" width="75" height="50" />

3 Bereitstellung von Beiträgen über Microformate in HTML-Seiten

Die primäre und empfohlene Inputquelle für den openTA-Newsdienst sind Newsfeeds der NTA-Mitgliedsinstitutionen.

Zukünftig sollen NTA-Institutionen auch die Möglichkeit haben, ihre Nachrichten über HTML-Seiten in ihrem Web-Auftritt in den openTA-Newsdienst einbinden zu lassen, wenn diese Seiten einen definierten, über ein Mikroformat [7] ausgezeichneten Aufbau haben.

Für den Newsdienst könnte die „CreativeWork“-Microdata-Klasse von schema.org verwendet werden, die auf rNews, einem Metadatenformat zur Beschreibung News-relevanter HTML-Seiten basiert [10]. Dabei würden für den openTA-Newsdienst nur die zum Atomformat passenden Metadaten verwendet. Eine Alternative ist das hAtom-Mikroformat [8], das auf News-Beiträge zugeschnitten und an das Atom-Feedformat angelehnt ist. hAtom liegt allerdings bisher nur in einer Entwurfsfassung vor, sodass Änderungen und Anpassungen nicht ausgeschlossen werden können.

Eine Entscheidung über die Realisierung dieser Option und das zu verwendende Format erfolgt erst, wenn seitens der NTA- Mitgliedsinstitutionen entsprechender Bedarf signalisiert wird!

4 Eingabe von Beiträgen im openTA-Portal über ein Webformular

Für NTA-Mitglieder, die aus technischen Gründen die oben beschriebenen Wege zur Bereitstellung von Informationen nicht nutzen können, könnte zusätzlich die Möglichkeit geschaffen, Newsbeiträge „per Hand“ im Mitgliederbereich des openTA-Portals zu erfassen. Dazu wird ein Formular zur Verfügung gestellt, das folgende Felder enthält:

Tab. 5: Vorgesehene Felder für ein Webformular für die Eingabe von Nachrichten

Feld	Anzahl Eingaben	Inhalt
Id	0	Eindeutiger Bezeichner, wird generiert
Titel	1	Text
Sprache	1	Auswahl: deutsch (Default) oder englisch
Untertitel	0-1	Text
Link	0-1	URL (Referenz auf Web-Seite zu dem Beitrag)
Zusammenfassung	0-1	Text
Inhalt	1	Text
Datum der Veröffentlichung	0-1	Datum (Vorschlagswert: aktuelles Tagesdatum)
Letzte Änderung	0	Datum und Uhrzeit (aktuell)
Autor	1-n	wird beschrieben durch die folgenden drei Felder; Der erste Autoreneintrag wird mit den im Portal bekannten Angaben zu dem Erfassenden vorbesetzt (änderbar)
Autor/Name	1	Name Person oder Institution
Autor/Email	0-1	Email-Adresse
Autor/URI	0-1	URI
Kategorie	1-n	Auswahl aus den vordefinierten Kategorien [3] und Freitext
Copyright	0-1	Text; die Creative Commons Lizenzen sollen als Auto Complete Texte bei der Eingabe angeboten werden.
Bild	0-1	URL oder Upload
WGS84-Koordinate	0-1	geografische Breite und Länge bezogen auf den Inhalt des Einzelbeitrags

Anzahl Eingaben: 0 = Ausgabefeld

1 = genau eine Eingabe (Pflicht)

0-1 = maximal eine Eingabe (optional)

1-n = ein oder mehrere Eingaben (mindestens eine Eingabe ist Pflicht)

0-n = mehrere Eingaben möglich (optional)

Zusätzlich zur Erfassung eines Beitrags soll die Möglichkeit geschaffen werden, einen über seine Id identifizierten Beitrag zu ändern oder zu löschen (nur durch den ursprünglichen Ersteller).

Die Funktion zur manuellen Erfassung von Beiträgen soll im Anschluss an die erfolgreiche Testphase mit Feed-basierten Nachrichten realisiert werden.

5 Nutzung des openTA-Newsdienstes

Die aggregierten Beiträge der NTA-Institutionen werden durch den openTA-Newsdienst auf verschiedenen Wegen zur Verfügung gestellt:

- Nachrichtenticker im openTA-Portal
- Konfigurierbarer Atom-Newsfeed
- *Recherche im News-Archiv (geplant)*
- Programmkomponente („Web-Widget“) zur Integration in beliebige Web-Seiten (bspw. Portalserver oder CMS)
- *E-Mail-Dienst (geplant)*

5.1 Nachrichtenticker im openTA-Portal

In das openTA-Portal wird ein Nachrichtenticker integriert, der die letzten (neuesten) vom openTA-Newsdienst aufgesammelten Beiträge anzeigt: <http://www.openta.net/news>. Zur Zeit werden in diesem Bereich Nachrichten, die nicht älter sind als drei Monate angezeigt. Diese Funktion basiert ebenso wie die in 5.2 und 5.3 beschriebenen Funktionalitäten auf

demselben technischen Grundgerüst. Es werden damit jedoch unterschiedliche Nutzungsszenarien realisiert.

5.2 Konfigurierbarer Atom-Newsfeed

Der openTA-Newsdienst erzeugt einen konfigurierbaren Atom-Newsfeed. Dazu wird im openTA-Portal öffentlich eine Konfigurationsseite angeboten, über die gewünschte Voreinstellungen vorgenommen werden können. Ergebnis der individuellen Konfiguration ist eine URL, über die der jeweilige Nutzer im Newsreader seiner Wahl den konfigurierten openTA-Newsfeed abonnieren kann. Ebenso kann jede NTA-Institution über diese URL den Feed beziehen und in die eigene Web-Seite einbauen.

Folgende Konfigurationsmöglichkeiten sind vorgesehen (keine Auswahl bei einem Parameter liefert alle Beiträge):

Tab. 6: Konfigurationsparameter für den openTA-Newsfeed

Parameter	Anzahl	Inhalt
Suchbegriff(e)	0-1	Freitext, verschiedene Suchbegriffe durch Leerzeichen trennbar (logische „oder“-Verknüpfung)
Organisation	0-n	Auswahl aus der Liste der NTA-Organisationen
Sprache	0-1	Auswahl: deutsch oder englisch
Kategorie	0-n	Auswahl aus den vordefinierten Kategorien [3]

Anzahl: 0-1 = maximal eine Option wählbar

0-n = Mehrfach-Auswahl möglich

Als zusätzliche Konfigurationsmöglichkeit für eine spätere Ausbaustufe ist die Vorgabe einer Bounding-Box über eine Karte zur regionalen Einschränkung angedacht. Zumindest im ersten Schritt dient eine solche Auswahl als Alternative zur Menüauswahl der die Nachrichten bereitstellenden NTA-Organisationen (s.o.). Eine Selektion über den per GeORSS definierten Raumbezug eines Beitrags wird geprüft, wenn genügend Beiträge mit der entsprechenden Information geliefert werden.

5.3 Recherche im News-Archiv

Die Newsbeiträge werden in openTA in einer Datenbank abgelegt, sodass auch auf alte Beiträge noch zugegriffen werden kann. Zur Recherche steht eine Service-Schnittstelle zur Verfügung, die über eine parametrisierbare URL angesprochen wird und das Ergebnis im Atom-Feedformat zurückliefert.

Syntax einer Anfrage:

```
http://service.openta.net/feeds/atompub/news?<Parameter1>=<Wert1>{&<Parameter2>=<Wert2>}
```

Die Recherche-URL wird durch Anhängen von Parameter-Werte-Paaren an eine Basis-URL gebildet. Es können folgende Parameter verwendet werden:

Tab. 7: Rechercheparameter für die Suche im openTA-News-Archiv

Parameter	Anzahl	Wert
q	0-1	Suchbegriff(e) für Volltextsuche in allen Textfeldern (ODER-Verknüpfung), Trennung durch Leerzeichen
org	0-n	NTA-Organisation (ODER-Verknüpfung)
lang	0-1	de oder en: Sprache deutsch oder englisch
c	0-n	Kategorie nach [3] (ODER-Verknüpfung); die Filterung nach einer übergeordneten Kategorie findet auch Beiträge, die nur mit einer dazu untergeordneten Kategorie verknüpft sind
from	0-1	Datum im Format yyyy-mm-dd: Beiträge mit dem Datum der letzten Änderung danach
until	0-1	Datum im Format yyyy-mm-dd: Beiträge mit dem Datum der letzten Änderung davor
indexOffset	0-1	Index (beginnend bei 0) ab dem die Einträge angezeigt werden sollen (für „Seitenblättern“)
count	0-1	Anzahl der Elemente die maximal zurückgeliefert werden. Nachfolgende Elemente können durch Anpassen des <i>indexOffset</i> -Parameters abgerufen werden.

Anzahl: 0-1 = maximal einmal zu verwenden

0-n = kann mehrfach verwendet werden

Eine Recherche kann auch interaktiv und formular-basiert über das openTA-Portal vorgenommen werden. Diese Funktion setzt auf der beschriebenen Service-Schnittstelle auf und kann dazu verwendet werden, eine parametrisierte URL zusammenzubauen.

5.4 Web-Widget

Das Nachrichten-Angebot von openTA kann auch schnell und unkompliziert in die eigene Webseite integriert werden. Zu diesem Zweck wird ein Web-Widget angeboten. Web-Widgets sind kleine, in sich abgeschlossene Elemente, welche einen bestimmten Bereich einer Webseite mit Fremdinformationen füllen.

Zum Einbinden des News-Widgets von openTA muss lediglich ein einfacher Code an die gewünschte Stelle zur Anzeige innerhalb der Webseite eingefügt werden. Um beispielsweise alle Nachrichten anzuzeigen, lautet der Codeschnipsel folgendermaßen:

```
<div data-widget="ota-news"></div>
<script type="text/javascript" src="//service.openta.net/widget/news-1.0.js"></script>
```

Darüber hinaus gibt es die Möglichkeit, die Auswahl der Nachrichten zu filtern und bestimmte Konfigurationsoptionen via HTML5-Datenattribute zu definieren:

Tab. 8: Datenattribute für das News-Widget

Datenattribut (data-...)	Bedeutung	Mögliche Werte
count	Konfiguration der Anzahl an Nachrichten, die pro Seite geladen werden	Ganzzahl > 0 (Standard: 5)
infinite	Konfiguration des „Infinite Scroll“-Features, welches neue Nachrichten beim Herunterscrollen direkt nachlädt, d. h. ohne Klick auf „mehr“-Link.	true false (Standard: false)
no-next	Konfiguration, ob Link zum Nachladen („mehr“-Link) angezeigt werden soll.	true false (Standard: false)
query	Filterung nach Suchtext	Beliebige Zeichenkette (Kodierung?)
itemlang	Filterung nach Sprache der Einträge	de en (Standard: leer)
fcats	Filterung nach formaler Kategorie	Kommaseparierte Liste von Ganzzahlen (ID der Kategorien, Standard: leer)
org	Filterung nach anliefernder Organisation	Kommaseparierte Liste von Ganzzahlen (ID der Organisationen, Standard: leer)

Für das Einbetten des Widgets, welches

- 10 Nachrichten pro Seite, (`data-count="10"`)
- mit Nachladen direkt beim Herunterscrollen, (`data-infinite="true"`)
- gefiltert nach dem Wort „Entwicklung“, (`data-query="Entwicklung"`)
- mit ausschließlich deutschen Nachrichten, (`data-itemlang="de"`)
- welche in der Kategorie „Projekt“ eingeordnet sind, (`data-fcats="8"`)
- und von ITAS angeliefert wurden (`data-org="12481"`)

anzzeigt, ergibt sich der Einbindungs-Code:

```
<div data-count="10" data-infinite="true" data-query="Entwicklung"
data-itemlang="de" data-fcats="8" data-org="12481"
data-widget="ota-news"></div>
<script type="text/javascript" src="//service.openta.net/widget/news-
1.0.js"></script>
```

Zur Vereinfachung der Einbettung (sowie Konfiguration) des News-Widgets sei die Verwendung des Code-Generators empfohlen. Hierzu muss einfach die News-Seite von openTA (<http://www.openta.net/news>) aufgerufen und mithilfe der Filtermöglichkeiten die Wunschauswahl zusammengestellt werden (die Ergebnisanzeige kann auch direkt betrachtet werden). Anschließend erzeugt der Klick auf „Widget selbst einbinden“ (vgl. Abb. 2) den passenden HTML-Code für die Einbettung.

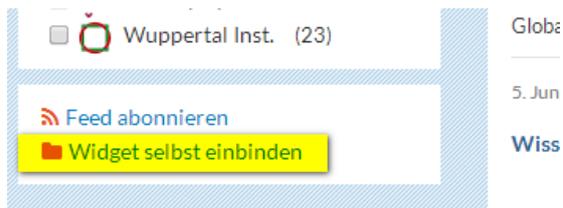


Abb. 2: openTA-News-Widget selbst einbinden

Die Nutzung des Generators hat darüber hinaus den Vorteil, dass dieser bei Software-Updates stets den aktuellen Stand bereitstellt. Bei Updates werden die älteren Versionen der Widgets jedoch weiter funktionieren, lediglich für neu hinzukommende Features kann eine Änderung des Einbettungscodes auf der Webseite nötig werden.

5.5 E-Mail-Dienst

Als zusätzliches Angebot ist ein Email-Dienst geplant. Über das openTA-Portal soll die Möglichkeit geschaffen werden, dass Nutzer sich für den Email-Dienst an- und abmelden. In regelmäßigen Abständen (einmal monatlich) erhalten die Abonnenten per Email eine Liste der seit der letzten Versendung neu eingesammelten News (Kurzbeschreibungen mit Link auf vollständigen Beitrag). Optional wird der Dienst konfigurierbar gestaltet mit denselben Parametern wie der News-Feed (Kap. 5.2).

Dieser Dienst wird in einer zweiten Ausbaustufe realisiert.

6 Literatur/Informationsquellen

1. Beschreibung Atom-Format:
<http://www.atomenabled.org/developers/syndication/index.php>
2. Beschreibung RSS 2.0 Format: <http://www.rssboard.org/rss-2-0>
3. openTA: Formale Kategorisierung, openTA-Dokument Nr. 1:
<http://www.openta.net/veroeffentlichungen-und-openta-dokumente>
4. RFC 1766: "Tags for the Identification of Languages", H. Alvestrand, March 1995,
<http://www.ietf.org/rfc/rfc1766.txt>
5. XML-Namensraum: [http://de.wikipedia.org/wiki/Namensraum_\(XML\)](http://de.wikipedia.org/wiki/Namensraum_(XML))
6. GeoRSS-simple: <http://georss.org/simple>
7. Mikroformate: <http://de.wikipedia.org/wiki/Mikroformate>
8. hAtom 0.1: <http://microformats.org/wiki/hatom>
9. MediaRSS: <http://www.rssboard.org/media-rss>
10. CreativeWork: <http://schema.org/CreativeWork>
11. Creative Commons Lizenzen: <http://creativecommons.org/licenses/>

7 Anhang A: Beispiel-Feed mit openTA-Kategorien im Atom-Format

Der folgende News-Feed im Atom-Format dient als Beispiel eines Feeds einer NTA-Organisation, der oben beschriebenen Vorgaben entspricht und daher von openTA problemlos verarbeitet werden könnte.

```
<feed xmlns="http://www.w3.org/2005/Atom">
  <id>DemoOrganisation1News</id>
  <title type="text">Demo-Organisation 1 News</title>
  <author>
    <name>Demo-Organisation 1</name>
    <email>info@example.org</email>
    <uri>http://www.example.org/1/</uri>
  </author>
  <link href="http://www.example.org/1/atomfeed.xml" rel="self" />
  <entry xml:lang="de">
    <id>123456789</id>
    <title type="text">Demo-Titel 1</title>
    <updated>2013-06-18T14:31:00.000Z</updated>
    <author>
      <name>Max Mustermann</name>
      <uri>http://www.example.org/1/max.mustermann/</uri>
    </author>
    <rights type="html">
      <a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">
        CC BY-NC-SA 3.0
      </a>
    </rights>
    <summary type="text">
      Lorem ipsum dolor sit amet, consectetur adipiscing elit.
    </summary>
    <content type="text">
      Lorem ipsum dolor sit amet, consectetur adipiscing elit.
      Aenean commodo ligula eget dolor. Aenean massa.
    </content>
    <link href="http://www.example.org/1/news/123456789"
rel="alternate" />
    <category term="veranst-ankuend"
scheme="http://service.openta.net/feeds/atompub/scheme/formal-
category"
label="Veranstaltung/Ankündigung" />
  </entry>
</feed>
```

8 Anhang B: Beispiel-Feed mit openTA-Kategorien im RSS-Format

Der folgende News-Feed im RSS-Format dient als Beispiel eines Feeds einer NTA-Organisation, der oben beschriebenen Vorgaben entspricht und daher von openTA problemlos verarbeitet werden könnte.

```
<rss version="2.0"
  xmlns:content="http://purl.org/rss/1.0/modules/content/">
  <channel>
    <title>Demo-News 1</title>
    <link>http://www.example.org/1</link>
    <language>de</language>
    <lastBuildDate>Thu, 20 Jun 2013 18:00:00 +0200</lastBuildDate>
    <item>
      <guid>http://www.example.org/1/news/123456789</guid>
      <title>Demo-Titel 1</title>
      <link>http://www.example.org/1/news/123456789</link>
      <description>
        Lorem ipsum dolor sit amet, consectetur adipiscing elit.
      </description>
      <content:encoded>
        <![CDATA[Lorem ipsum dolor sit amet, consectetur adipiscing
elit.
  Aenean commodo ligula eget dolor. Aenean massa.]]>
      </content:encoded>
      <category>openTA/Veranstaltung/Ankündigung</category>
      <pubDate>Thu, 20 Jun 2013 18:00:00 +0200</pubDate>
    </item>
  </channel>
</rss>
```